Analysis and chromosomal mapping of *Leishmania (Leishmania) amazonensis* amastigote expressed sequence tags

Luciana Girotto Gentil, Fernanda Lasakosvitsch, José Franco da Silveira, Márcia Regina Machado dos Santos, Clara Lúcia Barbiéri[+]

Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de São Paulo, Escola Paulista de Medicina, Rua Botucatu 862, 6º andar, 04023-062 São Paulo, SP, Brasil

The characterization of expressed sequence tags (ESTs) generated from a cDNA library of Leishmania (Leishmania) amazonensis *amastigotes is described. The sequencing of 93 clones generated new* L. (L.) amazonensis *ESTs from which 32% are not related to any other sequences in database and 68% presented significant similarities to known genes. The chromosome localization of some* L. (L.) amazonensis *ESTs was also determined in* L. (L.) amazonensis *and* L. (L.) major. *The characterization of these ESTs is suitable for the genome physical mapping, as well as for the identification of genes encoding cysteine proteinases implicated with protective immune responses in leishmaniasis.*

Key words: *Leishmania (Leishmania) amazonensis* - amastigotes - cDNA library

Sequencing of cDNA libraries to generate expressed sequence tags (ESTs) has been used for gene identification in *Plasmodium falciparum* (Chakrabarti et al. 1994), *Trypanosoma brucei* (El Sayed et al. 1995), *Trypanosoma cruzi* (Brandão et al. 1997, Verdún et al. 1998), and *Leishmania (L.) major* (Levick et al. 1996). The *L. (L.) major* genome project provided the sequence of the 36 chromosomes of the 32.8-megabase haploid genome of the parasite, predicting 911 RNA genes, 39 pseudogenes, and 8272 protein-coding genes (Ivens et al. 2005). In contrast, the genome of *L. (L.) amazonensis*, a causative agent of human cutaneous leishmaniasis in Brazil, is poorly studied and only 107 of its genes have been reported (http://ncbi.nlm.nih.gov). The availability of a cDNA library from amastigote forms of *Leishmania* can provide the characterization of genes implicated in parasite survival and host parasite interactions, opening perspectives for the development of new tools for disease control. The aim of the present work was the analysis of ESTs generated from a cDNA library of *L. (L.) amazonensis* amastigotes and their mapping in chromosomal bands.

In our laboratory we constructed a cDNA library from the amastigote form of *L. (L.) amazonensis*. It is important to emphasize that the amastigotes used for RNA extraction were derived from hamster foot lesions. This provides a measure of confidence of our cDNA library, since it was demonstrated that there is an increase in the magnitude of the transcript levels in *Leishmania* amastigotes proceeding from axenic cultures (Holzer et al. 2006). We obtained 40,000 clones and 100 were randomly selected and sequenced generating new *L. (L.) amazonensis* ESTs with an average insert size of 1.3 kb. These ESTs were sequenced from the 5' and 3' ends with T7 primer (forward) and SP6 primer (reverse) originating 147 sequences with an average size of 458 bp. These ESTs were compared to database sequences using the BLAST program. Sequence homologies identified by BLAST programs were considered statiscally significant with a Poisson *P* value of $\leq 10^{-4}$. Among 100 sequences, six matched with rRNA and one with hamster gene were excluded from further analysis. From a total of 93 EST sequences, 63 (68%) showed significant identity to gene sequences available on databases, whereas 30 ESTs (32%) did not have significant matches in databases and therefore were classified as undentified. Genes from other *Leishmania* species were matched to 64.5% of the ESTs, 2.1% matched to *L. (L.) amazonensis* sequences, and 1.1% matched to genes from other organisms. Considering both the identified and the undentified ESTs we could estimate the level of redundancy by use of SeqManII (DNAStar), a sequence analysis software. The overall diversity was 88.3% against 11.7% of redundancy. The most frequent genes on this analysis were those encoding the *L. (L.) mexicana* cysteine proteinase lmcpb2.8 (GenBank accession no. Z49962) and ribosomal phosphoprotein S6 (GenBank accession no. AF045457), which appeared three and two times. This level of redundancy was significantly lower than those observed in a non normalized library of *L. (L.) major* (Levick et al. 1996) and in other protozoan parasites (Chakrabarti et al. 1994, El Sayed et al. 1995, Brandão et al. 1997) whose redundancy was about 27%.

Further analysis of our data were performed by considering only nonredundant ESTs that showed significant match in the databases. The distribution of the *L. (L.) amazonensis* ESTs according to their putative cellular functions is shown in the Table. Most clones are related to protein synthesis (15%) that encode predominantly

TABLE

Functional classification of *Leishmania (L.) amazonensis* expressed sequence tags (ESTs)

Putative biological functions of ESTs	%
Protein synthesis	15.0
Transport facilitation	10.0
Cellular biogenesis	8.3
Metabolism	6.7
Protein destination	5.0
Intracellular transport	3.3
Signal transduction	3.3
Transcription	1.7
Hypotetical protein	41.7
Other	5.0

ribosomal proteins, corroborating data previously obtained in other protozoa (El Sayed et al. 1995, Verdún et al. 1998, Levick et al. 1996, Djikeng et al. 1998). Other clones are related to transport facilitation (10%), cellular biogenesis (8.3%), metabolism (6.7%), protein destination (5%), including two *Leishmania* cysteine proteinase genes, intracellular transport (3.3%), signal transduction (3.3%), and transcription (1.7%), whereas 41.7% were not identified to known proteins. This identification was based on the annotation present in the database GeneDB of *L. (L.) major* (www.gene-db.org/genedb/leish). Only two ESTs, *LlaAm*0013T7 and *LlaAm*0080 (GenBank accession nos. DY763238 and DY763159), showed identity to *L. (L.) amazonensis* genes deposited in database, translation elongation factor and Ca^{2+}-ATPase, respectively. This could be due to the scarcity of data relative to this *Leishmania* species. However, we identified several genes not yet described in *L. (L.) amazonensis* and one unidentified in trypanosomatids, the *LlaAm*0008SP6, a F1 ATP synthase (GenBank accession no. DY763167). We have also compared our results to previous studies employing mass spectrometry and microarray experiments for proteomic analysis of *L. (L.) mexicana* differentiation (Djikeng et al. 1998). Five ESTs identified in our studies were detected by these authors and two are exclusively expressed in amastigotes, the translation elongation factor-1 and the 40S ribosomal protein S12 (GenBank accession nos. DY763238 and DY763240, respectively). The other three ESTs were alpha-tubulin, ATPase beta subunit, and cysteine proteinase (GenBank accession nos. DY763245, DY763167, and DY763163, respectively). Recent data from microarray analysis of *L. (L.) major* and *L. (L.) mexicana* transcription are also in agreement with our results. Genes encoding cysteine proteinase and cytochrome b were predominantly transcribed in amastigote forms. Three ESTs identified in *L. (L.) amazonensis* (GenBank accession nos. DY763245, DY763151, and DY763244, respectively) correspond to alpha-tubulin, histone H2B and 60S ribosomal protein L13 transcripts detected almost exclusively in promastigotes, whereas one *L. (L.) amazonensis* EST (GenBank accession no. DY763267)

corresponds to amino acid permease only detected in metacyclic promastigotes (Saxena et al. 2003, Akopyants et al. 2004, Almeida et al. 2004, Holzer et al. 2006), confirming that transcription is constitutive among trypanosomatids. The significant number of *L. (L.) amazonensis* ESTs belonging to a hypothetical protein category described in the present work corroborates data obtained by microarray analysis of *L. (L.) mexicana* and demonstrates the importance to characterize genes related to specific functions in *Leishmania* (Saxena et al. 2003).

The identified ESTs and genes have been often used as chromosomal markers for determination of molecular karyotype as well as for construction of physical and genetic maps in trypanosomatids (Ivens et al. 1998, Melville et al. 2000, Porcile et al. 2003). The sequence analysis of our clones showed that some ESTs had identities with sequences previously annotated in *L. (L.) major* chromosomes (www.ncbi.nlm.nih.gov). In order to confirm the chromosomal polymorphism often observed in trypanosomatids (Henriksson et al. 1990) as well as to identify new chromosomal markers, we mapped the *L. (L.) amazonensis* ESTs in pulsed field gel electrophoresis (PFGE) karyotype of *L. (L.) amazonensis* and *L. (L.) major*. For this analysis, eight *L. (L.) amazonensis* ESTs were chosen: two (*LlaAm*0053 and *LlaAm*0108) presented identity with chromosome 14, two (*LlaAm*0044 and *LlaAm*0047) with chromosome 19, three (*LlaAm*0134, *LlaAm*0048, and *LlaAm*0087) with chromosomes 4, 21, and 27, respectively, and one (*LlaAm*0088) with chromosomes 21 and 35 from *L. (L.) major*. The *L. (L.) amazonensis* cysteine proteinase genes *Llacys*1 and *Llacys*2 previously cloned in our laboratory (Lasakosvitsch et al. 2003) and the 24Sα-rRNA were also used as probes and presented identity with *L. (L.) major* chromosomes 19, 8, and 27, respectively. A total of 25 chromosomal bands of *L. (L.) amazonensis*, ranging in size from 0.2 to ≥ 2.2 Mb, were resolved by PFGE, corroborating previous results obtained by Conte and Cano (2005). Some of bands showed variable ethidium bromide-staining intensities possibly due to co-migration of chromosomes of similar size (Figure). In control experiments using total DNA of *L. (L.) amazonensis* or *L. (L.) major* as a probe, all chromosomal bands were labeled by the probe, confirming their integrity and identity as chromosomes (data not shown). The hybridization of ESTs, cysteine proteinase genes, and 24Sα-rRNA with *L. (L.) amazonensis* chromosomal bands showed that nine of them mapped in a single band, whereas two EST (*LlaAm*0088) and the 24Sα-rRNA hybridized to two chromosomal bands with different sizes (Figure). The results obtained with seven markers (*LlaAm*0044, *LlaAm*0047, *LlaAm*0048, *LlaAm*0087, *LlaAm*0088, *Llacys*1 gene, and 24Sα-rRNA) were in agreement with data available in the *L. (L.) major* GeneDB. These genes are localized in chromosomes with different size, suggesting that they may represent homologue chromosomes with different size, as it has been observed in trypanosomatids (Henriksson et al. 1990, Santos et al. 1999).

Three genetic markers (*LlaAm*0088, *Llacys*2 gene, and 24Sα-rRNA) showed some polymorphism when we
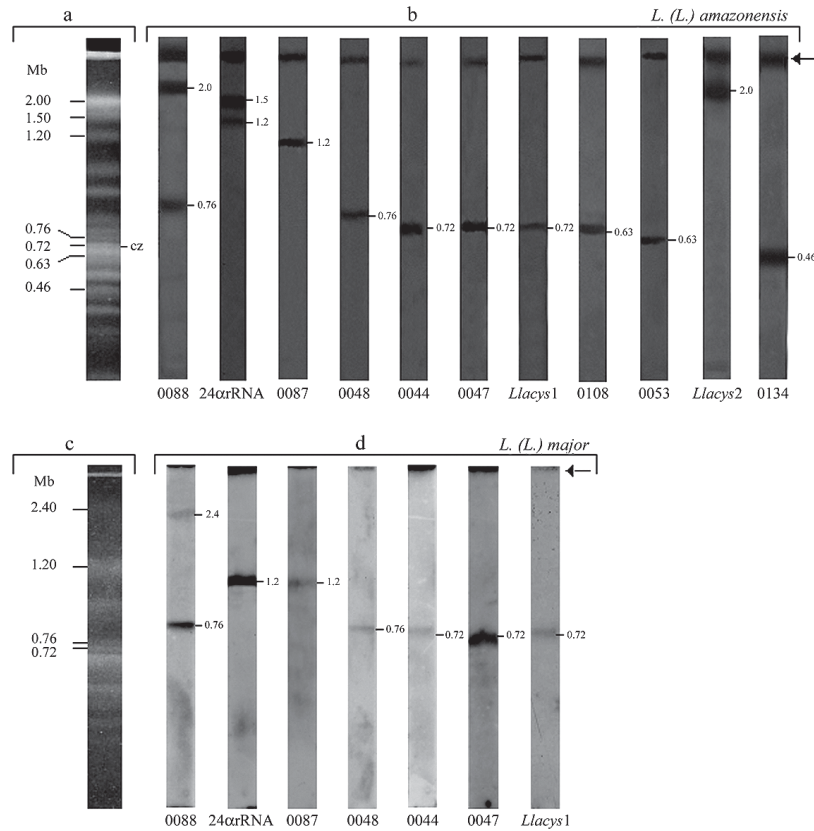
Mapping of markers on *Leishmania (L.) amazonensis* and *L. (L.) major* chromosomal bands. Chromosomes were separated by PFGE in the "Chef-DR®III System" (Bio-Rad). The electrophoresis for *L. (L.) amazonensis* chromosome separation was carried out in 1% agarose gels in 0.5 X TBE running buffer (45 mM Tris-borate, 1 mM EDTA, pH 8.3) at 14°C for 22 h with homogenous pulse time varying 60 s to 120 s at 6 V/cm. The running conditions for *L. (L.) major* were two phases of pulses with interpolation at 6 V/cm: phase 1, pulse time 50 s to 150 s (run time 16 h); phase 2, 150 s to 300 s (6 h). Gels were stained with 0.5 µg/ml ethidium bromide, photographed and transferred onto nylon filters. Chromosomal pattern of *L. (L.) amazonensis* (a) and *L. (L.) major* (c). Membranes were hybridized with radiolabeled probes at 42°C in 50% formamide overnight and washed in high stringency conditions; b: hybridization of *L. (L.) amazonensis* chromosomal bands with *L. (L.) amazonensis* markers, eight ESTs (*LlaAm*0088, *LlaAm*0087, *LlaAm*0048, *LlaAm*0044, *LlaAm*0047, *LlaAm*00108, *LlaAm*0053, and *LlaAm*00134), *Llacys*1 and *Llacys2* genes and 24Sα-rRNA; d: hybridization of *L. (L.) major* chromosomal bands with five EST markers (*LlaAm*0088, *LlaAm*0087, *LlaAm*0048, *LlaAm*0044, and *LlaAm*0047), *Llacys*1 gene and 24Sα-rRNA. The sizes (Mb) are indicated; cz: compression zone; the arrow indicates the front of the gel.

compared *L. (L.) amazonensis* and *L. (L.) major* karyotypes (Figure). The *LlaAm*0088 hybridized to two bands, 760 kb and 2000 kb, of *L. (L.) amazonensis*, suggesting a deletion of 400 kb in the homologue in *L. (L.) major*. An unequivocal evidence of chromosomal rearrangement could also be observed with the 24Sα-rRNA which mapped on a chromosomal band of 1200 kb in both species. In addition, a band of 1500 kb contains the 24Sα-rRNA in *L. (L.) amazonensis*. It is also possible to observe that the *Llacys2* gene mapped on a 565 kb band in *L. (L.) major* and on a single band of 2000 kb in *L. (L.) amazonensis*. Among the nine markers that hybridized to a single band in *L. (L.) amazonensis*, five ESTs did not show any known identity, two were identified as a ribosomal protein (*LlaAm*0048) and ATP/ADP translocase (*LlaAm*0044), and two as cysteine proteinase genes (*Llacys*1 and *Llacys2*). These clones will be useful as chromosome-specific markers to identify polymorphism among *L. (L.) amazonensis* and other *Leish-*

mania species. The ESTs *LlaAm*0044, *LlaAm*0047, and *Llacys*1 hybridized in a same chromosomal band of 720 kb in *L. (L.) amazonensis* and *L. (L.) major*. Analysis of the chromosome 19 sequence from *L. (L.) major* (www.genedb.org/genedb/leish) showed that the clones *LlaAm*0044, *LlaAm*0047, and *Llacys*1 are located on 50 kb, 190 kb, and 620 kb of this chromosome, respectively. These results suggest that these genes are located in the same chromosome in *L. (L.) amazonensis*, indicating that they belong to a linkage group and can be used as markers for a specific chromosome. Similarly, the *LlaAm*0053 and *LlaAm*00108 ESTs mapped on a chromosomal band of the same size (630 kb) in both species and they could also be used as specific chromosomal markers. In contrast, the markers *LlaAm*0088, *Llacys2* gene, and 24Sα-rRNA showed different hybridization patterns in the two *Leishmania* species, strengthening the chromosomal polymorphism previously reported among *Leishmania* species from the Old and New World (Britto et al. 1998).

Cloning of cysteine proteinase genes represents another interesting approach for use of the cDNA library described in the present work. Cysteine proteinases have been considered virulence factors in *Leishmania* (Denise et al. 2003) and targets for the design of vaccines (Rafati et al. 2001). Previous studies from our laboratory demonstrated that a cysteine proteinase of 30 kDa, predominantly expressed in amastigote forms of *L. (L.) amazonensis*, can confer protective immune responses against homologous infection in BALB/c mice (Beyrodt et al. 1997). Cloning of *L. (L.) amazonensis Llacys1* and *Llacys2* genes has led to the development of new schedules of immunization using both recombinant cysteine proteinases as well as *Llacys1* and *Llacys2* genes.

REFERENCES

Akopyants NS, Matlib RS, Bukanova EN, Smeds MR, Brownstein BH, Stormo GD, Beverley SM 2004. Expression profiling using random genomic DNA microarrays identifies differentially expressed genes associated with three major developmental stages of the protozoan parasite *Leishmania major*. *Mol Biochem Parasitol 36*: 71-86.

Almeida R, Gilmartin BJ, McCann SH, Norrish A, Ivens AC, Lawson D, Levick, MP, Smith DF, Dyall SD, Vetrie D, Freeman TC, Coulson RM, Sampaio I, Schneider H, Blackwell JM 2004. Expression profiling of the *Leishmania* life cycle: cDNA arrays identify developmentally regulated genes present but not annotated in the genome. *Mol Biochem Parasitol 136*: 87-100.

Beyrodt CGP, Pinto AR, Freymüller E, Barbiéri CL 1997. Characterization of an antigen from *Leishmania amazonensis* amastigotes able to elicit protective immune responses in a murine model. *Infect Immun 65*: 2052-2059.

Brandão A, Urmenyi T, Rondinelli E, González A, de Miranda AB, Degrave W 1997. Identification of transcribed sequences (ESTs) in the *Trypanosoma cruzi* genome project. *Mem Inst Oswaldo Cruz 92*: 863-866.

Britto C, Ravel C, Bastien P, Blaineau C, Pages M, Dedet JP, Wincker P 1998. Conserved linkage groups associated with large-scale chromosomal rearrangements between Old World and New World *Leishmania* genomes. *Gene 222*: 107-117.

Chakrabarti D, Reddy GR, Dame JB, Almira EC, Laipis PJ, Ferl RJ, Yang TP, Rowe TC, Schuster SM 1994. Analysis of expressed sequence tags from *Plasmodium falciparum*. *Mol Biochem Parasitol 66*: 97-104.

Conte FF, Cano MI 2005. Genomic organization of telomeric and subtelomeric sequences of *Leishmania (Leishmania) amazonensis*. *Int J Parasitol 35*: 1435-1443.

Denise H, McNeil K, Brooks DR, Alexander J, Coombs GH, Mottram JC 2003. Expression of multiple CPB genes encoding cysteine proteases is required for *Leishmania mexicana* virulence in vivo. *Infect Immun 71*: 3190-3195.

Djikeng A, Agufa C, Donelson JE, Majiwa PA 1998. Generation of expressed sequence tags as physical landmarks in the genome of *Trypanosoma brucei*. *Gene 221*: 93-106.

El Sayed NMA, Alarcon CM, Beck JC, Sheffield VC, Donelson JE 1995. cDNA expressed sequence tags of *Trypanosoma brucei rhodesiense* provide new insights into the biology of the parasite. *Mol Biochem Parasitol 73*: 75-90.

Henriksson J, Aslund L, Macina RA, Franke de Cazzulo BM, Cazzulo JJ, Frasch AC, Pettersson U 1990. Chromosomal localization of seven cloned antigen genes provides evidence of diploidy and further demonstration of karyotype variability in *Trypanosoma cruzi*. *Mol Biochem Parasitol 42*: 213-223.

Holzer TR, McMaster WR, Forney JD 2006. Expression profiling by whole-genome interspecies microarray hybridization reveals differential gene expression in procyclic promastigotes, lesion-derived amastigotes, and axenic amastigotes in *Leishmania mexicana*. *Mol Biochem Parasitol 146*: 198-218.

Ivens AC, Lewis SM, Bagherzadeh A, Zhang L, Chan HM, Smith DF 1998. A physical map of the *Leishmania major* Friedlin genome. *Genome Res 8*: 135-45.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream, MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchettin G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RMR, Cronin A, Cruz AK, Davies RM, Gaudenzi JD, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie L, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Müller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'Neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabbinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schäfer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart JD, Barrell B, Myler PJ 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science 309*: 436-442.

Lasakosvitsch F, Gentil LG, Santos MRM, Silveira JF, Barbiéri CL 2003. Cloning and characterisation of a cysteine proteinase gene expressed in amastigotes of *Leishmania (L.) amazonensis*. *Int J Parasitol 33*: 445-454.

Levick MP, Blackwell JM, Connor V, Coulson RMR, Miles A, Smith HE, Wan KL, Ajioka JW 1996. An expressed sequence tag analysis of a full-length, spliced-leader cDNA library from *Leishmania major* promastigotes. *Mol Biochem Parasitol 76*: 345-348.

Melville SE, Leech V, Navarro M, Cross GA 2000. The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* stock 427. *Mol Biochem Parasitol 111*: 261-73.

Nugent PG, Karsani SA, Wait R, Tempero J, Smith, DF 2004. Proteomic analysis of *Leishmania mexicana* differentiation. *Mol Biochem Parasitol 136*: 51-62.

Porcile PE, Santos MR, Souza RT, Verbisck NV, Brandão A, Urmenyi T, Silva R, Rondinelli E, Lorenzi H, Levin MJ, Degrave W, Franco da Silveira J 2003. A refined molecular karyotype for the reference strain of the *Trypanosoma cruzi* genome project (clone CL Brener) by assignment of chromosome markers. *Gene 308*: 53-65.

Rafati S, Salmanian AH, Taheri T, Vafa M, Fasel N 2001. A protective cocktail vaccine against murine cutaneous leishma-

niasis with DNA encoding cysteine proteinases of *Leishmania major. Vaccine 19*: 3369-3375.

Santos MRM, Lorenzi H, Porcile P, Carmo MS, Schijman A, Brandão A, Araya JE, Gomes HB, Chiurillo MA, Ramirez JL, Degrave WM, Levin MJ, Silveira JF 1999. Physical mapping of a 670-kb region of chromosomes XVI and XVII from the human protozoan parasite *Trypanosoma cruzi* encompassing the genes for two immunodominant antigens. *Genome Res 9*: 1268-1276.

Saxena A, Worthey EA, Yan S, Leland A, Stuart KD, Myler PJ 2003. Evaluation of differential gene expression in *Leishmania major* Friedlin procyclics and metacyclics using DNA microarray analysis. *Mol Biochem Parasitol 129*: 103-114.

Verdún RE, Di Paolo N, Urmenyi TP, Rondinelli E, Frasch AC, Sánchez DO 1998. Gene discovery through expressed sequence tag sequencing in *Trypanosoma cruzi. Infect Immun 66*: 5393-5398.